Available online at www.sciencedirect.com

SCIENCE DIRECT®

J. Math. Anal. Appl. 298 (2004) 374–397

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.elsevier.com/locate/jmaa

A unifying local–semilocal convergence analysis and applications for two-point Newton-like methods in Banach space

Ioannis K. Argyros

Cameron University, Department of Mathematical Sciences, Lawton, OK 73505, USA

Received 15 September 2003

Submitted by W.L. Wendland

Abstract

We provide a local as well as a semilocal convergence analysis for two-point Newton-like methods in a Banach space setting under very general Lipschitz type conditions. Our equation contains a Fréchet differentiable operator F and another operator G whose differentiability is not assumed. Using more precise majorizing sequences than before we provide sufficient convergence conditions for Newton-like methods to a locally unique solution of equation $F(x) + G(x) = 0$. In the semilocal case we show under weaker conditions that our error estimates on the distances involved are finer and the information on the location of the solution at least as precise as in earlier results. In the local case a larger radius of convergence is obtained. Several numerical examples are provided to show that our results compare favorably with earlier ones. As a special case we show that the famous Newton–Kantorovich hypothesis is weakened under the same hypotheses as the ones contained in the Newton–Kantorovich theorem.

Keywords: Newton-like method; Banach space; Majorizing sequence; Fréchet-derivative; Newton–Kantorovich method/hypothesis; Radius of convergence; Banach lemma on invertible operators

E-mail address: ioannisa@cameron.edu.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2004.04.008

1. Introduction

In this study we are concerned with the problem of approximating a locally unique solution x^* of the nonlinear equation

$$F(x) + G(x) = 0, \tag{1}$$

where F, G are operators defined on a closed ball $\bar{U}(w, R)$ centered at point w and of radius $R \geqslant 0$, which is a subset of a Banach space X with values in a Banach space Y. F is Fréchet-differentiable on $\bar{U}(w, R)$, while the differentiability of the operator G is not assumed.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations [4,7,16]. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time-invariant system is driven by the equation $\dot{x} = Q(x)$ (for some suitable operator Q), where x is the state. Then the equilibrium states are determined by solving Eq. (1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative—when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

We use the two-point Newton method

$$\begin{aligned} &y_{-1}, y_0 \in \bar{U}(w, R), \\ &y_{n+1} = y_n - A(y_{n-1}, y_n)^{-1}\big[F(y_n) + G(y_n)\big] \quad (n \geqslant 0) \end{aligned} \tag{2}$$

to generate a sequence converging to x^*. Here $A(x, y) \in L(X, Y)$, the space of bounded linear operators from X into Y for each fixed $x, y \in \bar{U}(w, R)$. We provide a local as well as a semilocal convergence analysis for method (2) under very general Lipschitz-type hypotheses (see (25), (26)).

Our new idea is to use center-Lipschitz conditions instead of Lipschitz conditions for the upper bounds on the inverses of the linear operators involved. It turns out that this way we obtain more precise majorizing sequences. Moreover, despite the fact that our conditions are more general than related ones already in the literature [1–26], we can provide weaker sufficient convergence conditions, and finer error bounds on the distances involved.

We note that our analysis is also useful in particular in the numerical solution of problems appearing in visco-elasticity [4,7]. However we leave the details to the motivated reader. Finally we mention that our approach compares favorably with the classical and elegant work of J.W. Schmidt on the Secant method (see [20–22] and our Example 2).

Several applications are provided: e.g., in the semilocal case we show that the famous Newton–Kantorovich hypothesis (for its simplicity and transparency, see (5)) is weakened (see (20)), whereas in the local case we can provide a larger convergence radius using the same information (see (134) and (135)).

2. Semilocal convergence analysis of method (2)

Part A. Motivation

Deuflhard and Heindl [13] have proved the following affine invariant form of the Newton–Kantorovich theorem [16] which is the motivation for this study.

Theorem 1. *Let $F : D \subseteq X \to Y$ be a Fréchet-differentiable operator on an open convex set D. Suppose that $d_0 \in D$ is such that $F'(d_0)^{-1}$ exists and*

$$\left\| F'(d_0)^{-1} F(d_0) \right\| \leqslant \eta, \tag{3}$$

$$\left\| F'(d_0)^{-1} \left[F'(x) - F'(y) \right] \right\| \leqslant \gamma_1 \|x - y\| \quad \textit{for all } x, y \in D \textit{ and } \gamma_1 > 0, \tag{4}$$

$$h = 2\gamma_1 \eta \leqslant 1, \tag{5}$$

$$\bar{U}(d_0, d^1) = \left\{ x \in X \mid \|x - d_0\| \leqslant d^1 \right\} \subseteq D, \tag{6}$$

where

$$d^1 = \frac{1 - \sqrt{1 - h}}{\gamma_1}. \tag{7}$$

Then, sequence $\{d_n\}$ $(n \geqslant 0)$ generated by Newton's method

$$d_{n+1} = d_n - F'(d_n)^{-1} F(d_n) \quad (n \geqslant 0), \tag{8}$$

is well defined, remains in $\bar{U}(d_0, d^1)$ for all $n \geqslant 0$ and converges to a unique solution d^* of equation

$$F(d) = 0 \tag{9}$$

in $\bar{U}(d_0, d^1) \cup (D \cap U(d_0, d^2))$, where

$$d^2 = \frac{1 + \sqrt{1 - h}}{\gamma_1}. \tag{10}$$

Moreover the following error bounds hold:

$$\|d_{n+1} - d_n\| \leqslant \bar{d}_{n+1} - \bar{d}_n, \tag{11}$$

$$\|d_n - d^*\| \leqslant d^1 - \bar{d}_n, \quad d^1 = \lim_{n \to \infty} \bar{d}_n, \tag{12}$$

where sequence $\{\bar{d}_n\}$ $(n \geqslant 0)$ is given by

$$\bar{d}_0 = 0, \quad \bar{d}_1 = \eta, \quad \bar{d}_{n+2} = \bar{d}_{n+1} + \frac{\gamma_1 (\bar{d}_{n+1} - \bar{d}_n)^2}{2(1 - \gamma_1 \bar{d}_{n+1})}. \tag{13}$$

Condition (5) is the famous Newton–Kantorovich hypothesis which is the essential sufficient convergence condition for the semilocal convergence of Newton's method (8). However Newton's method may converge to a solution of Eq. (9) even when (5) is violated.

Example 1. Let $X = Y = \mathbf{R}$, $d_0 = 1$, $D = [p, 2 - p]$, $p \in [0, 1/2)$, and define F on D by

$$F(d) = d^3 - p. \tag{14}$$

Using (3), (4) and (14) we get

$$\eta = \frac{1}{3}(1 - p), \qquad \gamma_1 = 2(2 - p), \tag{15}$$

which imply

$$h = \frac{4}{3}(1 - p)(2 - p) > 1 \quad \text{for all } p \in \left[0, \frac{1}{2}\right). \tag{16}$$

That is, there is no guarantee that method (8) converges since (5) is violated. However one can find values of p in $[0, 1/2)$ such that method (8) converges.

For example, if $p = 0.48$, then using (8) we find $d^* = \sqrt[3]{0.48}$. Hence, we wonder if (5) can be weakened. Hypothesis (5) is used to show that majorizing sequence $\{\bar{d}_n\}$ is monotonically increasing and converges to d^1. We have noticed that sequence $\{\bar{\bar{d}}_n\}$ $(n \geqslant 0)$ given by

$$\bar{\bar{d}}_n = 0, \quad \bar{\bar{d}}_1 = \eta, \quad \bar{\bar{d}}_{n+2} = \bar{\bar{d}}_{n+1} + \frac{\gamma_1(\bar{\bar{d}}_{n+1} - \bar{\bar{d}}_n)^2}{2(1 - \gamma_0\bar{\bar{d}}_{n+1})} \tag{17}$$

is also a more precise majorizing sequence for Newton's method (8) than (13), where γ_0 is the center-Lipschitz constant such that

$$\left\| F'(d_0)^{-1}\left[F'(x) - F'(d_0)\right]\right\| \leqslant \gamma_0\|x - d_0\| \quad \text{for all } x \in D. \tag{18}$$

In general the inequality

$$\gamma_0 \leqslant \gamma_1 \tag{19}$$

holds. Note also that in practice finding constant γ_1 requires the computation of γ_0. Hence no additional computational effort is required to compute (γ_0, γ_1) instead of γ_1. As it is shown in a more general setting in what follows (see Application 2) in this case (5) can be replaced by

$$h_1 = (\gamma_0 + \gamma_1)\eta \leqslant 1 \tag{20}$$

(see also (72)). By comparing (5) and (20) we get $h_1 \geqslant h/2$. Moreover note that

$$h \leqslant 1 \quad \Rightarrow \quad h_1 \leqslant 1 \tag{21}$$

but not vice versa unless if $\gamma_0 = \gamma_1$. Furthermore as it is shown in a more general setting for all $n \geqslant 0$ $(\gamma_0 < \gamma_1)$,

$$\|d_{n+1} - d_n\| \leqslant \bar{\bar{d}}_{n+1} - \bar{\bar{d}}_n < \bar{d}_{n+1} - \bar{d}_n, \tag{22}$$

$$\|d_n - d^*\| \leqslant d^3 - \bar{\bar{d}}_n \leqslant d_1 - \bar{d}_n, \tag{23}$$

and

$$d^3 = \lim_{n\to\infty} \bar{\bar{d}}_n \leqslant d_1 \tag{24}$$

(see Remark 5). Hence we also obtain finer error bounds and at least as precise information on the location of the solution d^* as in Theorem 1. Returning back to Example 1, since $\gamma_0 = 3 - p$ we find that (20) holds if $p \in [(5 - \sqrt{13})/3, 1/2)$, which improves Theorem 1.

Part B. Main results

In order for us to show that these observations hold in a more general setting we first need to introduce the following assumptions.

Let $R \geqslant 0$ be given. Assume there exist $v, w \in X$ such that $A(v,w)^{-1} \in L(Y,X)$, and for any $x, y, z \in \bar{U}(w,r) \subseteq \bar{U}(w,R)$, $t \in [0,1]$, the following hold:

$$\left\| A(v,w)^{-1}\left[A(x,y) - A(v,w)\right] \right\| \leqslant h_0\left(\|x-v\|, \|y-w\|\right) + a \tag{25}$$

and

$$\begin{aligned} &\left\| A(v,w)^{-1}\left\{\left[F'\left(y+t(z-y)\right) - A(x,y)\right](z-y) + G(z) - G(y)\right\} \right\| \\ &\quad \leqslant \left[h_1\left(\|y-w\| + t\|z-y\|\right) - h_2\left(\|y-w\|\right) + h_3\left(\|z-x\|\right) + b\right]\|z-y\|, \end{aligned} \tag{26}$$

where $h_0(r,s)$, $h_1(r+\bar{r}) - h_2(r)$ $(\bar{r} \geqslant 0)$, $h_2(r)$, $h_3(r)$ are monotonically increasing functions for all r, s on $[0,R]^2$, $[0,R]^2$, $[0,R]$, $[0,R]$, respectively, with $h_0(0,0) = h_1(0) = h_2(0) = h_3(0) = 0$, and the constants a, b satisfy $a \geqslant 0$, $b \geqslant 0$. Given y_{-1}, y_0, v, w in X, define parameters c_{-1}, c, c_1 by

$$\|y_{-1} - v\| \leqslant c_{-1}, \quad \|y_{-1} - y_0\| \leqslant c, \quad \|v - w\| \leqslant c_1. \tag{27}$$

Remark 1. Conditions similar to (21)–(26) but less flexible were considered by Chen and Yamamoto in [11] in the special case when $A(x,y) = A(x)$ for all $x, y \in \bar{U}(w,R)$ $(A(x) \in L(X,Y))$ (see also Theorem 4). Operator $A(x)$ is intended there to be an approximation to the Fréchet-derivative $F'(x)$ of F. However we also want the choice of operator A to be more flexible, and be related to the difference $G(z) - G(y)$ for all $y, z \in \bar{U}(w,R)$. It has already been shown in special cases in [3,4,10] that this way the ratio of convergence for method (2) is improved (see also Application 1). Note also that if we choose

$$\begin{aligned} &A(x,y) = F'(x), \quad G(x) = 0, \quad w = d_0, \\ &h_0(r,r) = \gamma_0 r, \quad h_1(r) = h_2(r) = \gamma_1 r, \quad h_3(r) = 0 \end{aligned} \tag{28}$$

for all $x, y \in \bar{U}(w,R)$, $r \in [0,R]$, and $a = b = 0$, then conditions (25) and (26) reduce to (18) and (4), respectively. Other choices of operators, functions and constants appearing in (25) and (26) can be found in the applications that follow.

With the above choices, we show the following result on majorizing sequences for method (2).

Lemma 1. *Assume that there exist parameters $\eta \geqslant 0$, $a \geqslant 0$, $b \geqslant 0$, $c_{-1} \geqslant 0$, $c \geqslant 0$, $\delta \in [0,2)$, $r_0 \in [0,R]$ such that*

$$2\left[\int_0^1 h_1(r_0 + \theta\eta)\,d\theta - h_2(r_0) + b + h_3(c+\eta)\right] + \left[a + h_0(c + c_{-1}, \eta + r_0)\right]\delta \leqslant \delta, \tag{29}$$

$$\frac{2\eta}{2-\delta} + r_0 + c \leqslant R, \tag{30}$$

$$h_0\left[\frac{1-\left(\frac{\delta}{2}\right)^{n+1}}{1-\frac{\delta}{2}}\eta + c + c_{-1}, \frac{1-\left(\frac{\delta}{2}\right)^{n+2}}{1-\frac{\delta}{2}}\eta + r_0\right] + a < 1, \tag{31}$$

and

$$\begin{aligned} &2\int_0^1 h_1\left[\frac{1-\left(\frac{\delta}{2}\right)^{n+1}}{1-\frac{\delta}{2}}\eta + \theta\left(\frac{\delta}{2}\right)^{n+1}\eta + r_0\right]d\theta - 2h_2\left[\frac{1-\left(\frac{\delta}{2}\right)^{n+1}}{1-\frac{\delta}{2}}\eta + r_0\right] \\ &\quad + 2h_3\left[\left(\frac{\delta}{2}\right)^n\left(1+\frac{\delta}{2}\right)\eta\right] + \delta h_0\left[\frac{1-\left(\frac{\delta}{2}\right)^{n+1}}{1-\frac{\delta}{2}}\eta + c + c_{-1}, \frac{1-\left(\frac{\delta}{2}\right)^{n+2}}{1-\frac{\delta}{2}}\eta + r_0\right] \\ &\leqslant \delta \end{aligned} \tag{32}$$

for all $n \geqslant 0$. Then, iteration $\{t_n\}$ $(n \geqslant -1)$ given by

$$\begin{aligned} &t_{-1} = r_0, \quad t_0 = c + r_0, \quad t_1 = c + r_0 + \eta, \\ &t_{n+2} = t_{n+1} + \frac{\{\int_0^1 [h_1(t_n - t_0 + r_0 + \theta(t_{n+1} - t_n)) - h_2(t_n - t_0 + r_0) + b]\,d\theta + h_3(t_{n+1} - t_{n-1})\}(t_{n+1} - t_n)}{1 - a - h_0(t_n - t_{-1} + c_{-1}, t_{n+1} - t_0 + r_0)} \end{aligned} \tag{33}$$

is monotonically increasing, bounded above by

$$t^{**} = \frac{2\eta}{2-\delta} + r_0 + c, \tag{34}$$

and converges to some t^* such that

$$0 \leqslant t^* \leqslant t^{**} \leqslant R. \tag{35}$$

Moreover the following error bounds hold for all $n \geqslant 0$:

$$0 \leqslant t_{n+2} - t_{n+1} \leqslant \frac{\delta}{2}(t_{n+1} - t_n) \leqslant \left(\frac{\delta}{2}\right)^{n+1}\eta. \tag{36}$$

Proof. We must show

$$\begin{aligned} &2\left\{\int_0^1 \left[h_1\big(t_k - t_0 + r_0 + \theta(t_{k+1} - t_k)\big) - h_2(t_k - t_0 + r_0) + b\right]d\theta \right. \\ &\quad \left. + h_3(t_{k+1} - t_{k-1})\right\} + \delta\left[a + h_0(t_k - t_{-1} + c_{-1}, t_{k+1} - t_0 + r_0)\right] \leqslant \delta, \end{aligned} \tag{37}$$

$$0 \leqslant t_{k+1} - t_k, \tag{38}$$

and

$$h_0(t_k - t_{-1} + c_{-1}, t_{k+1} - t_0 + r_0) + a < 1 \tag{39}$$

for all $k \geqslant 0$.

Estimate (36) can then follow from (37)–(39) and (33).

Using induction on the integer $k \geqslant 0$, first for $k = 0$ in (37)–(39) we must show

$$2\left[\int_0^1 h_1(r_0+\theta\eta)\,d\theta - h_2(r_0) + b + h_3(c+\eta)\right] + \delta\bigl[a + h_0(c+c_{-1}, \eta + r_0)\bigr] \leqslant \delta,$$
$$0 \leqslant t_1 - t_0,$$
$$h_0(c+c_{-1}, \eta + r_0) + a < 1,$$

which hold by (29) and the definition of t_1.

By (33) we get

$$0 \leqslant t_2 - t_1 \leqslant \frac{\delta}{2}(t_1 - t_0).$$

Assume that (37)–(39) hold for all $k \leqslant n+1$. Using (37)–(39) we obtain in turn

$$\begin{aligned}
&2\left\{\int_0^1 \bigl[h_1\bigl(t_{k+1} - t_0 + r_0 + \theta(t_{k+2} - t_{k+1})\bigr) - h_2(t_{k+1} - t_0 + r_0) + b\bigr]\,d\theta \right.\\
&\qquad \left. + h_3(t_{k+2} - t_k)\right\} + \delta\bigl[a + h_0(t_{k+1} - t_{-1} + c_{-1}, t_{k+2} - t_0 + r_0)\bigr]\\
&\quad \leqslant 2\left\{\int_0^1 h_1\left[\left(\frac{1-(\frac{\delta}{2})^{k+1}}{1-\frac{\delta}{2}} + \theta\left(\frac{\delta}{2}\right)^{k+1}\right)\eta + r_0\right] - h_2\left[\frac{1-(\frac{\delta}{2})^{k+1}}{1-\frac{\delta}{2}}\eta + r_0\right]\right.\\
&\qquad \left. + b + h_3\left[\left(\frac{\delta}{2}\right)^{k+1}\eta + \left(\frac{\delta}{2}\right)^{k}\eta\right]\right\}\\
&\qquad + \delta\left[a + h_0\left(\frac{1-(\frac{\delta}{2})^{k+1}}{1-\frac{\delta}{2}}\eta + c + c_{-1}, \frac{1-(\frac{\delta}{2})^{k+2}}{1-\frac{\delta}{2}}\eta + r_0\right)\right]\\
&\quad \leqslant \delta
\end{aligned}$$

by (29) and (32). Hence we showed (37) holds for $k = n+2$. Moreover, we must show

$$t_k \leqslant t^{**}, \tag{40}$$
$$t_{-1} = r_0 \leqslant t^{**}, \quad t_0 = r_0 + c \leqslant t^{**}, \quad t_1 = c + r_0 + \eta \leqslant t^{**},$$
$$t_2 \leqslant c + r_0 + \eta + \frac{\delta}{2}\eta = \frac{2+\delta}{2}\eta + r_0 + c \leqslant t^{**}.$$

Assume that (40) holds for all $k \leqslant n+1$. It follows from (33), (37)–(39) that

$$\begin{aligned} t_{k+2} &\leqslant t_{k+1} + \frac{\delta}{2}(t_{k+1} - t_k) \leqslant t_k + \frac{\delta}{2}(t_k - t_{k-1}) + \frac{\delta}{2}(t_{k+1} - t_k) \\ &\leqslant \cdots \leqslant c + r_0 + \eta + \frac{\delta}{2}\eta + \left(\frac{\delta}{2}\right)^2 \eta + \cdots + \left(\frac{\delta}{2}\right)^{k+1} \eta \\ &= \frac{1 - \left(\frac{\delta}{2}\right)^{k+2}}{1 - \frac{\delta}{2}}\eta + r_0 + c \leqslant \frac{2\eta}{2 - \delta} + r_0 + c = t^{**}. \end{aligned} \tag{41}$$

Hence, sequence $\{t_n\}$ $(n \geqslant -1)$ is bounded above by t^{**}. Inequality (39) holds for $k = n + 2$ by (30) and (31). Moreover (38) holds for $k = n + 2$ by (41) and since (37) and (39) also hold for $k = n + 2$. Furthermore, sequence $\{t_n\}$ $(n \geqslant 0)$ is monotonically increasing by (38) and as such it converges to some t^* satisfying (35).

That completes the proof of Lemma 1. □

We provide the main result on the semilocal convergence of method (2) using majorizing sequence (33).

Theorem 2. *Assume that hypotheses of Lemma* 1 *hold, and there exist*

$$y_{-1} \in \bar{U}(w, R), \qquad y_0 \in \bar{U}(w, r_0) \tag{42}$$

such that

$$\left\| A(y_{-1}, y_0)^{-1}\left[F(y_0) + G(y_0)\right]\right\| \leqslant \eta. \tag{43}$$

Then, sequence $\{y_n\}$ $(n \geqslant -1)$ *generated by Newton-like method* (2) *is well defined, remains in* $\bar{U}(w, t^*)$ *for all* $n \geqslant -1$*, and converges to a solution* x^* *of equation* $F(x) + G(x) = 0$*. Moreover the following error bounds hold for all* $n \geqslant -1$:

$$\|y_{n+1} - y_n\| \leqslant t_{n+1} - t_n \tag{44}$$

and

$$\|y_n - x^*\| \leqslant t^* - t_n. \tag{45}$$

Furthermore the solution x^* *is unique in* $\bar{U}(w, t^*)$ *if*

$$\int_0^1 h_1\big((1+t)t^*\big)\,dt - h_2(t^*) + h_3(t^*) + h_0(t^* + c_1, t^*) + a + b < 1, \tag{46}$$

and in $\bar{U}(w, R_0)$ *for* $R_0 \in (t^*, R]$ *if*

$$\int_0^1 h_1(t^* + tR_0)\,dt - h_2(t^*) + h_3(R_0) + h_0(t^* + c_1, t^*) + a + b < 1, \tag{47}$$

provided that $x_{-1} = v$ *and* $x_0 = w$.

Proof. We first show estimate (44), and $y_n \in \bar{U}(w, t^*)$ for all $n \geqslant -1$. For $n = -1, 0$, (44) follows from (27), (33) and (43). Suppose (44) holds for all $n = 0, 1, \ldots, k + 1$; this implies in particular (using (27), (42))

$$\begin{aligned}\|y_{k+1}-w\| &\leqslant \|y_{k+1}-y_k\|+\|y_k-y_{k-1}\|+\cdots+\|y_1-y_0\|+\|y_0-w\|\\ &\leqslant (t_{k+1}-t_k)+(t_k-t_{k-1})+\cdots+(t_1-t_0)+r_0=t_{k+1}-t_0+r_0\\ &\leqslant t_{k+1}\leqslant t^*.\end{aligned}$$

That is, $y_{k+1}\in \bar{U}(w,t^*)$.

We show (44) holds for $n=k+2$. By (25) and (33) we obtain for all $x,y\in \bar{U}(w,t^*)$,

$$\big\|A(v,w)^{-1}\big[A(x,y)-A(v,w)\big]\big\|\leqslant h_0\big(\|x-v\|,\|y-w\|\big)+a. \tag{48}$$

In particular for $x=y_k$ and $y=y_{k+1}$ we get using (25), (27),

$$\begin{aligned}&\big\|A(v,w)^{-1}\big[A(y_k,y_{k+1})-A(v,w)\big]\big\|\leqslant h_0\big(\|y_k-v\|,\|y_{k+1}-w\|\big)+a\\ &\quad\leqslant h_0\big(\|y_k-y_{-1}\|+\|y_{-1}-v\|,\|y_{k+1}-x_0\|+\|y_0-w\|\big)+a\\ &\quad\leqslant h_0(t_k-t_{-1}+c_{-1},t_{k+1}-t_0+r_0)+a\\ &\quad\leqslant h_0\left[\frac{1-\left(\frac{\delta}{2}\right)^k}{1-\frac{\delta}{2}}\eta+c+c_{-1},\ \frac{1-\left(\frac{\delta}{2}\right)^{k+1}}{1-\frac{\delta}{2}}\eta+r_0\right]+a<1\quad\text{(by (31))}.\end{aligned}\tag{49}$$

It follows from (49) and the Banach lemma on invertible operators [16] that $A(y_k,y_{k+1})^{-1}$ exists, and

$$\big\|A(y_k,y_{k+1})^{-1}A(v,w)\big\|\leqslant\big[1-a-h_0(t_k-t_{-1}+c_{-1},t_{k+1}-t_0+r_0)\big]^{-1}. \tag{50}$$

Using (2), (26), (33), (50) we obtain in turn

$$\begin{aligned}\|y_{k+2}-y_{k+1}\| &= \big\|A(y_k,y_{k+1})^{-1}\big[F(y_{k+1})+G(y_{k+1})\big]\big\|\\ &= \big\|A(y_k,y_{k+1})^{-1}\big[F(y_{k+1})+G(y_{k+1})-A(y_{k-1},y_k)(y_{k+1}-y_k)\\ &\qquad -F(y_k)-G(y_k)\big]\big\|\\ &\leqslant \big\|A(y_k,y_{k+1})^{-1}A(v,w)\big\|\,\big\|A(v,w)^{-1}\big[F(y_{k+1})-F(y_k)\\ &\qquad -A(y_{k-1},y_k)(y_{k+1}-y_k)+G(y_{k+1})-G(y_k)\big]\big\|\\ &\leqslant \tfrac{\{\int_0^1[h_1(\|y_k-w\|+t\|y_{k+1}-y_k\|)-h_2(\|y_k-w\|)+b]\,dt+h_3(\|y_{k+1}-y_{k-1}\|)\}\|y_{k+1}-y_k\|}{1-a-h_0(t_k-t_{-1}+c_{-1},t_{k+1}-t_0+r_0)}\\ &\leqslant \tfrac{\{\int_0^1[h_1(t_k-t_0+r_0+t(t_{k+1}-t_k))-h_2(t_k-t_0+r_0)+b]\,dt+h_3(t_{k+1}-t_{k-1})\}(t_{k+1}-t_k)}{1-a-h_0(t_k-t_{-1}+c_{-1},t_{k+1}-t_0+r_0)}\\ &= t_{k+2}-t_{k+1},\end{aligned}\tag{51}$$

which shows (44) for all $n\geqslant 0$.

Note also that

$$\begin{aligned}\|y_{k+2}-y_{k+1}\| &\leqslant \|y_{k+2}-y_{k+1}\|+\|y_{k+1}-z\|\leqslant t_{k+2}-t_{k+1}+t_{k+1}-t_0+r_0\\ &= t_{k+2}-t_0+r_0\leqslant t_{k+2}\leqslant t^*.\end{aligned}\tag{52}$$

That is, $y_{k+2}\in \bar{U}(z,t^*)$.

It follows from (44) that $\{y_n\}$ $(n\geqslant -1)$ is a Cauchy sequence in a Banach space X, and as such it converges to some $x^*\in X\in \bar{U}(w,t^*)$ (since $\bar{U}(w,t^*)$ is a closed set). By letting $k\to\infty$ in (51) we obtain $F(x^*)+G(x^*)=0$. Estimate (45) follows from (44) by using standard majorization techniques [8,16].

To show uniqueness in $\bar{U}(w,t^*)$, let y^* be a solution of Eq. (1) in $\bar{U}(w,t^*)$. We define Newton-like iteration $\{x_n\}$ $(n \geqslant -1)$ by

$$x_{-1} = v, \qquad x_0 = w,$$
$$x_{n+1} = x_n - A(x_{n-1}, x_n)^{-1}\big[F(x_n) + G(x_n)\big] \quad (n \geqslant 0). \tag{53}$$

Iteration $\{x_n\}$ $(n \geqslant -1)$ is a special case of $\{y_n\}$ $(n \geqslant -1)$. Hence, we have

$$\|x_{k+1} - x_k\| \leqslant \bar{t}_{k+1} - \bar{t}_k, \qquad \lim_{n\to\infty} x_n = x^*,$$

and

$$\|x^* - x_k\| \leqslant t^* - \bar{t}_k, \qquad \lim \bar{t}_k = t^*, \tag{54}$$

where $\{\bar{t}_n\}$ is $\{t_n\}$ $(n \geqslant -1)$ for $r_0 = 0$.

We shall show

$$\|y^* - x_k\| \leqslant t^* - \bar{t}_k. \tag{55}$$

For $k = 0$, (54) holds since $y^* \in \bar{U}(w,t^*)$. Suppose (54) holds for all $n \leqslant k$. Then as in (51) we obtain the identity

$$\begin{aligned} y^* - x_{k+1} &= y^* - x_k + A(x_{k-1}, x_k)^{-1}\big(F(x_k) + G(x_k)\big) \\ &\quad - A(x_{k-1}, x_k)^{-1}\big(F(y^*) + G(y^*)\big) \\ &= \big[A(x_{k-1}, x_k)^{-1}A(x_{-1}, x_0)\big]A(x_{-1}, x_0)^{-1}\big[F(x_k) - F(y^*) \\ &\quad - A(x_{k-1}, x_k)(y^* - x_k) + G(x_k) - F(y^*)\big]. \end{aligned} \tag{56}$$

Using (56) we obtain in turn

$$\begin{aligned} \|y^* - x_{k+1}\| &\leqslant \frac{[\int_0^1 h_1(\|x_k - x_0\| + t\|y^* - x_k\|)\,dt - h_2(\|x_k - x_0\|) + h_3(\|y^* - x_{k-1}\|) + b]\|y^* - x_k\|}{1 - a - h_0(\|x_{k-1} - x_{-1}\|, \|x_n - x_0\|)} \\ &\leqslant \frac{[\int_0^1 h_1((1+t)t^*)\,dt - h_2(t^*) + h_3(t^*) + b]}{1 - a - h_0(t^* + c_1, t^*)}\|y^* - x_k\| \\ &< \|y^* - x_k\| \leqslant t^* - \bar{t}_k \to 0 \quad \text{as } k \to \infty. \end{aligned} \tag{57}$$

That is, $x^* = y^*$.

If $y^* \in \bar{U}(x_0, R_0)$ then as in (57) we get

$$\begin{aligned} \|y^* - x_{k+1}\| &\leqslant \frac{[\int_0^1 h_1(t^* + tR_0)\,dt - h_2(t^*) + h_3(R_0) + b]}{1 - a - h_0(t^* + c_1, t^*)}\|y^* - x_k\| \\ &< \|y^* - x_k\|. \end{aligned} \tag{58}$$

Hence, again we get $x^* = y*$.

That completes the proof of Theorem 2. □

Remark 2. Conditions (31), (32) can be replaced by the stronger but easier to check

$$h_0\left[\frac{2\eta}{2-\delta} + c + c_{-1}, \frac{2\eta}{2-\delta} + r_0\right] + a \leqslant 1 \tag{59}$$

and

$$2\int_0^1 h_1\left[\frac{2\eta}{2-\delta}+\theta\frac{\delta}{2}+r_0\right]d\theta-2h_2\left[\frac{2\eta}{2-\delta}+r_0\right]$$
$$+2h_3\left[\left(1+\frac{\delta}{2}\right)\eta\right]+\delta h_0\left[\frac{2\eta}{2-\delta}+c+c_{-1},\frac{2\eta}{2-\delta}+r_0\right]$$
$$\leqslant\delta, \tag{60}$$

respectively. Note also that conditions (29)–(32), (59), (60) are of the Newton–Kantorovich-type hypotheses (see also (5)) which are always present in the study of Newton-like methods [4,8,11,12,24,26].

Application 1. Let us consider some special choices of operator A, functions h_i, $i=0,1,2,3$, parameters a, b and points v, w.

Define

$$A(x,y)=F'(y)+[x,y;G], \tag{61}$$

$$v=y_{-1},\qquad w=y_0, \tag{62}$$

and set

$$r_0=0, \tag{63}$$

where F', $[\cdot,\cdot;G]$ denote the Fréchet-derivative of F and the divided difference of order one for operator G, respectively [4,8,10]. Hence, we consider Newton-like method (2) in the form

$$y_{n+1}=y_n-\bigl(F'(y_n)+[y_{n-1},y_n;G]\bigr)^{-1}\bigl(F(y_n)+G(y_n)\bigr)\quad(n\geqslant 0). \tag{64}$$

The method was studied in [3,4,10]. It is shown to be of order $(1+\sqrt{5})/2\approx 1.618\ldots$ (same as the order of Chord), but higher than the order of

$$z_{n+1}=z_n-F'(z_n)^{-1}\bigl(F(z_n)+G(z_n)\bigr)\quad(n\geqslant 0) \tag{65}$$

and

$$w_{n+1}=w_n-A(w_n)^{-1}\bigl(F(w_n)+G(w_n)\bigr)\quad(n\geqslant 0), \tag{66}$$

where $A(\cdot)$ is an operator approximating F' (see, e.g., [4,8,11,12,21]). Assume

$$\bigl\|A(y_{-1},y_0)^{-1}\bigl[F'(y)-F'(y_0)\bigr]\bigr\|\leqslant\gamma_2\|y-y_0\|, \tag{67}$$

$$\bigl\|A(y_{-1},y_0)^{-1}\bigl[F'(x)-F'(y)\bigr]\bigr\|\leqslant\gamma_3\|x-y\|, \tag{68}$$

$$\bigl\|A(y_{-1},y_0)^{-1}\bigl([x,y;G]-[y_{-1},y_0;G]\bigr)\bigr\|\leqslant\gamma_4\bigl(\|x-y_{-1}\|+\|y-y_0\|\bigr), \tag{69}$$

and

$$\bigl\|A(y_{-1},y_0)^{-1}\bigl([x,y;G]-[z,x;G]\bigr)\bigr\|\leqslant\gamma_5\|z-y\| \tag{70}$$

for some nonnegative parameters γ_i, $i=2,3,4,5$, and all $x,y\in\bar{U}(y_0,r)\subseteq\bar{U}(y_0,R)$.

Then we can define

$$\begin{aligned}&a=b=0,\quad h_1=h_2,\quad h_1(q)=\gamma_3 q,\quad h_3(q)=\gamma_5 q,\quad \text{and}\\ &h_0(q_1,q_2)=\gamma_4 q_1+(\gamma_2+\gamma_4)q_2.\end{aligned}\tag{71}$$

If the hypotheses of Theorem 2 hold for the above choices, the conclusions follow.

Note that conditions (67)–(70) are weaker than the corresponding ones in [10, pp. 48–49], [3,4]. Indeed, conditions

$$\left\|F'(x)-F'(y)\right\|\leqslant\gamma_6\|x-y\|,\quad \left\|A(x,y)^{-1}\right\|\leqslant\gamma_7,\quad \left\|[x,y,z;G]\right\|\leqslant\gamma_8,$$

and

$$\left\|[x,y;G]-[z,w;G]\right\|\leqslant\gamma_9\big(\|x-z\|+\|y-w\|\big)$$

for all $x,y,z,w\in\bar{U}(y_0,r)$ are used there instead of (67)–(70), where $[x,y,z;G]$ denotes a second order divided difference of G at (x,y,z), and γ_i, $i=6,7,8,9$, are non-negative parameters.

Let us provide an example for this case.

Example 2. Let $X=Y=(\mathbf{R}^2,\|\cdot\|_\infty)$. Consider the system

$$\begin{aligned}&3x^2y+y^2-1+|x-1|=0,\\ &x^4+xy^3-1+|y|=0.\end{aligned}$$

Set $\|x\|_\infty=\|(x',x'')\|_\infty=\max\{|x'|,|x''|\}$, $F=(F_1,F_2)$, and $G=(G_1,G_2)$. For $x=(x',x'')\in\mathbf{R}^2$ we take $F_1(x',x'')=3(x')^2x''+(x'')^2-1$, $F_2(x',x'')=(x')^4+x'(x'')^3-1$, $G_1(x',x'')=|x'-1|$, $G_2(x',x'')=|x''|$. We shall take $[x,y;G]\in M_{2\times2}(\mathbf{R})$ as

$$[x,y;G]_{i,1}=\frac{G_i(y',y'')-G_i(x',y'')}{y'-x'},\quad [x,y;G]_{i,2}=\frac{G_i(x',y'')-G_i(x',x'')}{y''-x''},\quad i=1,2.$$

Using method (65) with $z_0=(1,0)$ we obtain

n	$z_n^{(1)}$	$z_n^{(2)}$	$\|z_n-z_{n-1}\|$
0	1	0	
1	1	0.333333333333333	3.333E−1
2	0.906550218340611	0.354002911208151	9.344E−2
3	0.885328400663412	0.338027276361322	2.122E−2
4	0.891329556832800	0.326613976593566	1.141E−2
5	0.895238815463844	0.326406852843625	3.909E−3
6	0.895154671372635	0.327730334045043	1.323E−3
7	0.894673743471137	0.327979154372032	4.809E−4
8	0.894598908977448	0.327865059348755	1.140E−4
9	0.894643228355865	0.327815039208286	5.002E−5
10	0.894659993615645	0.327819889264891	1.676E−5
11	0.894657640195329	0.327826728208560	6.838E−6
12	0.894655219565091	0.327827351826856	2.420E−6
13	0.894655074977661	0.327826643198819	7.086E−7
...			
39	0.894655373334687	0.327826511746298	5.149E−19

Using the method of chord (i.e., (66) with $A(w_n) = [w_{n-1}, w_1; G]$) with $w_0 = (5, 5)$, $w_{-1} = (1, 0)$, we obtain

n	$w_n^{(1)}$	$w_n^{(2)}$	$\|w_n - w_{n-1}\|$
0	5	5	
1	1	0	5.000E+00
2	0.989800874210782	0.012627489072365	1.262E−02
3	0.921814765493287	0.307939916152262	2.953E−01
4	0.900073765669214	0.325927010697792	2.174E−02
5	0.894939851625105	0.327725437396226	5.133E−03
6	0.894658420586013	0.327825363500783	2.814E−04
7	0.894655375077418	0.327826521051833	3.045E−04
8	0.894655373334698	0.327826521746293	1.742E−09
9	0.894655373334687	0.327826521746298	1.076E−14
10	0.894655373334687	0.327826521746298	5.421E−20

Using our method (64) with $y_0 = (5, 5)$, $y_{-1} = (1, 0)$, we obtain

n	$y_n^{(1)}$	$y_n^{(2)}$	$\|y_n - y_{n-1}\|$
0	5	5	
1	1	0	5
2	0.909090909090909	0.363636363636364	3.636E−01
3	0.894886945874111	0.329098638203090	3.453E−02
4	0.894655531991499	0.327827544745569	1.271E−03
5	0.894655373334793	0.327826521746906	1.022E−06
6	0.894655373334687	0.327826521746298	6.089E−13
7	0.894655373334687	0.327826421746298	2.710E−20

We did not verify the hypotheses of Theorem 3 for the above starting points. However, it is clear that the hypotheses of Theorem 3 are satisfied for all three methods for starting points closer to the solution

$$x^* = (0.894655373334687, 3.27826421746298)$$

chosen from the lists of the tables displayed above.

Hence method (2) (i.e., method (64) in this case) converges faster than (65) suggested in Chen and Yamamoto [11], Zabrejko and Nguen [26] in this case and the method of chord [21,22].

In the application that follows we show that the famous Newton–Kantorovich hypothesis (see (5)) is weakened under the same hypotheses/information [16,18,19,23].

Application 2. Returning back to Remark 1 and (28), iteration (2) reduces to the famous Newton–Kantorovich method (8).

Condition (29) reduces to

$$h_\delta = (\gamma_1 + \delta\gamma_0)\eta \leqslant \delta. \tag{72}$$

Case 1. Let us restrict $\delta \in [0, 1]$. Hypothesis (32) now becomes

$$2\int_0^1 \gamma_1\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{k+1}\right)+\theta\left(\frac{\delta}{2}\right)^{k+1}\eta\right]d\theta$$
$$-2\gamma_1\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{k+1}\right)\right]+\delta\gamma_0\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{k+1}\right)\right]$$
$$\leqslant 2\gamma_1\int_0^1 \theta\eta\,d\theta+\delta\gamma_0\eta$$

or

$$\left[\frac{\gamma_0\delta^2}{2-\delta}-\gamma_1\right]\left[1-\left(\frac{\delta}{2}\right)^{k+1}\right]\leqslant 1,$$

which is true for all $k \geqslant 0$ by the choice of δ. Furthermore (31) gives

$$\frac{2\gamma_0\eta}{2-\delta}\leqslant 2\gamma_0\eta<1.$$

Hence in this case conditions (29), (31) and (32) reduce only to (72) provided $\delta \in [0,1]$. Condition (72) for say $\delta = 1$ reduces to (20).

Case 2. It follows from Case 1 that (29), (31) and (32) reduce to (72),

$$\frac{2\gamma_0\eta}{2-\delta}\leqslant 1 \tag{73}$$

and

$$\frac{\gamma_0\delta^2}{2-\delta}\leqslant \gamma_1, \tag{74}$$

respectively, provided $\delta \in [0,2)$.

Case 3. It turns out that the range for δ can be extended (see also Example 3). Introduce conditions

$$\gamma_0\eta \leqslant 1-\frac{1}{2}\delta \quad \text{for } \delta \in [\delta_0, 2),$$

where

$$\delta_0=\frac{-b+\sqrt{b^2+8b}}{2},\quad b=\frac{\gamma_1}{\gamma_0},\quad \text{and}\quad \gamma_0\neq 0.$$

Indeed the proof of Theorem 2 goes through if instead we show the weaker condition

$$\gamma_1\left(\frac{\delta}{2}\right)^{k+1}+\frac{2\gamma_0\delta}{2-\delta}\left[1-\left(\frac{\delta}{2}\right)^{k+2}\right]\leqslant \delta,$$

or

$$\left(b-\frac{b\delta}{2}-\frac{\delta^2}{2}\right)\left(\frac{\delta}{2}\right)^{k+1}\leqslant 0,$$

or

$$\delta \geqslant \delta_0,$$

which is true by the choice of δ_0.

Example 3. Returning back to Example 1 but using Case 3 we can do better. Indeed, choose

$$p = p_0 = 0.4505 < \frac{5-\sqrt{13}}{3} = 0.464816242\ldots.$$

Then we get

$$\eta = 0.183166\ldots, \quad \gamma_0 = 2.5495, \quad \gamma_1 = 3.099, \quad \text{and} \quad \delta_0 = 1.0656867.$$

Choose $\delta = \delta_0$. Then we get

$$\gamma_0\eta = 0.466983415 < 1 - \frac{\delta_0}{2} = 0.46715665.$$

That is the interval $[(5-\sqrt{13})/2, 1/2)$ can be extended to at least $[p_0, 1/2)$.

In the example that follows we show that γ_1/γ_0 can be arbitrarily large. Indeed

Example 4. Let $X = Y = \mathbf{R}$, $d_0 = 0$ and define functions F, G on R by

$$F(x) = c_0x + c_1 + c_2 \sin e^{c_3x}, \qquad G(x) = 0, \tag{75}$$

where c_i, $i = 0, 1, 2, 3$, are given parameters. Using (75) it can easily be seen that for c_3 large and c_2 sufficiently small γ_1/γ_0 can be arbitrarily large.

Part C. Specialization to one-step methods

In order to compare with earlier results, we consider the case when $x = y$ and $v = w$ (single step methods). We can then prove along the same lines to Lemma 1 and Theorem 2, respectively, the following results by assuming that there exists $w \in X$ such that $A(w)^{-1} \in L(Y, X)$, for any $x, y \in \bar{U}(w, r) \subseteq \bar{U}(w, R)$, $t \in [0, 1]$,

$$\left\|A(w)^{-1}\left[A(x) - A(w)\right]\right\| \leqslant g_0\left(\|x - w\|\right) + \alpha \tag{76}$$

and

$$\begin{aligned} &\left\|A(w)^{-1}\left\{\left[F(x + t(y - x)) - A(x)\right](y - x) + G(y) - G(x)\right\}\right\| \\ &\quad \leqslant \left[g_1\left(\|x - w\| + t\|y - x\|\right) - g_2\left(\|x - w\|\right) + g_3(r) + \beta\right]\|y - x\|, \end{aligned} \tag{77}$$

where $g_0, g_1, g_2, g_3, \alpha, \beta$ are as h_0 (one variable), h_1, h_2, h_3, a, b, respectively.

Then we can show the following result on majorizing sequences.

Lemma 2. *Assume that there exist $\eta \geqslant 0$, $\alpha \geqslant 0$, $\beta \geqslant 0$, $\delta \in [0, 2)$, $r_0 \in [0, R]$ such that*

$$\bar{h}_\delta = 2\left[\int_0^1 g_1(r_0+\theta\eta)\,d\theta - g_2(r_0) + g_3(r_0+\eta) + \beta\right] + \delta\left[\alpha + g_0(r_0+\eta)\right] \leqslant \delta, \tag{78}$$

$$\frac{2\eta}{2-\delta} + r_0 \leqslant R, \tag{79}$$

$$g_0\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] + \alpha < 1, \tag{80}$$

$$\begin{aligned} &2\int_0^1 g_1\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right) + r_0 + \theta\left(\frac{\delta}{2}\right)^{n+1}\eta\right] d\theta \\ &\quad -2g_2\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] + 2g_3\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] \\ &\quad + \delta g_0\left[\frac{2\eta}{2-\delta}\left(1-\left(\frac{\delta}{2}\right)^{n+1}\right) + r_0\right] \\ &\leqslant \delta \end{aligned} \tag{81}$$

for all $n \geqslant 0$. Then, iteration $\{s_n\}$ $(n \geqslant 0)$ given by

$$s_0 = r_0, \qquad s_1 = r_0 + \eta,$$
$$s_{n+2} = s_{n+1} + \frac{\int_0^1 \{g_1(s_n+\theta(s_{n+1}-s_n)) - g_2(s_n) + \beta\}\,d\theta\,(s_{n+1}-s_n) + \int_{s_n}^{s_{n+1}} g_3(\theta)\,d\theta}{1-\alpha-g_0(s_{n+1})} \tag{82}$$

is monotonically increasing, bounded above by

$$s^{**} = \frac{2\eta}{2-\delta} + r_0, \tag{83}$$

and converges to some s^* such that

$$0 \leqslant s^* \leqslant s^{**}. \tag{84}$$

Moreover the following error bounds hold for all $n \geqslant 0$:

$$0 \leqslant s_{n+2} - s_{n+1} \leqslant \frac{\delta}{2}(s_{n+1}-s_n) \leqslant \left(\frac{\delta}{2}\right)^{n+1}\eta. \tag{85}$$

Theorem 3. *Assume that hypotheses of Lemma 2 hold and there exists $y_0 \in \bar{U}(w, r_0)$ such that*

$$\left\|A(y_0)^{-1}\left[F(y_0) + G(y_0)\right]\right\| \leqslant \eta. \tag{86}$$

Then, sequence $\{w_n\}$ $(n \geqslant 0)$ generated by Newton-like method (66) is well defined, remains in $\bar{U}(w, s^*)$ for all $n \geqslant 0$, and converges to a solution x^* of equation $F(x) + G(x) = 0$. Moreover the following error bounds hold for all $n \geqslant 0$:

$$\|w_{n+1} - w_n\| \leqslant s_{n+1} - s_n \tag{87}$$

and

$$\|w_n - x^*\| \leqslant s^* - s_n. \tag{88}$$

Furthermore the solution x^* is unique in $\bar{U}(w, s^*)$ if

$$\int_0^1 \left[g_1(s^* + \theta s^*) - g_2(s^*)\right] d\theta + g_3(s^*) + g_0(s^*) + \alpha + \beta < 1, \tag{89}$$

or in $\bar{U}(w, R_0)$ if $s^* < R_0 \leqslant R$, and

$$\int_0^1 \left[g_1(s^* + \theta R_0) - g_2(s^*)\right] d\theta + g_3(s^* + R_0) + g_0(s^*) + \alpha + \beta < 1, \tag{90}$$

provided that $w_0 = w$.

We state the relevant results due to Chen and Yamamoto [11, p. 40]. We assume that $A(w)^{-1}$ exists, and for any $x, y \in \bar{U}(w, r) \subseteq \bar{U}(z, R)$,

$$0 < \left\|A(w)^{-1}\big(F(w) + G(w)\big)\right\| \leqslant \bar{\eta}, \tag{91}$$

$$\left\|A(w)^{-1}\big(A(x) - A(w)\big)\right\| \leqslant \bar{g}_0\big(\|x - w\|\big) + \bar{\alpha}, \tag{92}$$

$$\begin{aligned}&\left\|A(w)^{-1}\left[F'\big(x + t(y - x)\big) - A(x)\right]\right\| \\ &\quad \leqslant \bar{g}_1\big(\|x - w\|\big) + t\big(\|y - x\|\big) - \bar{g}_0\big(\|x - w\|\big) + \bar{\beta}, \quad t \in [0, 1],\end{aligned} \tag{93}$$

$$\left\|A(w)^{-1}\left[G(x) - G(y)\right]\right\| \leqslant g_3(r)\|x - y\|, \tag{94}$$

where $\bar{g}_0, \bar{g}_1, \bar{\alpha}, \bar{\beta}$ are as g_0, g_1, α, β, respectively, but $\bar{g}_0$ is also differentiable with $\bar{g}_0'(r) > 0$, $r \in [0, R]$, and $\bar{\alpha} + \bar{\beta} < 1$.

As in [11], set

$$\varphi(r) = \bar{\eta} - r + \int_0^r \bar{g}_1(t)\, dt, \qquad \psi(r) = \int_0^r g_3(t)\, dt, \tag{95}$$

$$\chi(r) = \phi(r) + \psi(r) + (\bar{\alpha} + \bar{\beta})r. \tag{96}$$

Denote the minimal value of $\chi(r)$ on $[0, R]$ by χ^*, and the minimal point by r^*. If $\chi(R) \leqslant 0$, denote the unique zero of χ by $r_0^* \in (0, r^*]$. Define scalar sequence $\{r_n\}$ $(n \geqslant 0)$ by

$$r_0 \in [0, R], \qquad r_{n+1} = r_n + \frac{u(r_n)}{g(r_n)} \quad (n \geqslant 0), \tag{97}$$

where

$$u(r) = \chi(r) - x^* \tag{98}$$

and

$$g(r) = 1 - \bar{g}_0(r) - \bar{\alpha}. \tag{99}$$

With the above notation they showed

Theorem 4 [11, p. 40]. *Suppose $\chi(R) \leqslant 0$. Then Eq.* (1) *has a solution $x^* \in \bar{U}(w, r_0^*)$, which is unique in*

$$\tilde{U} = \begin{cases} \bar{U}(w, R) & \text{if } \chi(R) = 0 \text{ or } \psi(R) = 0, \text{ and } r_0^* < R, \\ U(w, R) & \text{if } \chi(R) = 0 \text{ and } r_0^* < R. \end{cases} \tag{100}$$

Let

$$D^* = \bar{U}_{r \in [0, r^*)} \left\{ y \in \bar{U}(w, r) \;\middle|\; \left\| A(y)^{-1}\left[F(y) + G(y)\right]\right\| \leqslant \frac{u(r)}{g(r)} \right\}. \tag{101}$$

Then, for any $y_0 \in D$, sequence $\{y_n\}$ $(n \geqslant 0)$ generated by Newton-like method (66) *is well defined, remains in $\bar{U}(w, r^*)$ and satisfies*

$$\|y_{n+1} - y_n\| \leqslant r_{n+1} - r_n \tag{102}$$

and

$$\|y_n - x^*\| \leqslant r^* - r_n \tag{103}$$

provided that r_0 is chosen as in (97) *so that $r_0 \in Ry_0$, where for $y \in D^*$,*

$$R_y = \left\{ r \in [0, r^*) \;\middle|\; \left\| A(y)^{-1}\left(F(y) + G(y)\right)\right\| \leqslant \frac{u(r)}{y(r)},\ \|y - z\| \leqslant r \right\}. \tag{104}$$

Remark 3. (a) Hypothesis on $\bar{g}_0$ is stronger than the corresponding one on g_0.

(b) Iteration (97) converges to r^* (even if $r_0 = 0$) not r_0^*.

(c) Choices of y_{-1}, y_0 other than the ones in Theorems 2, 3 can be given by (101) and (102).

Remark 4. The conclusions of Theorem 4 hold (i.e., the results in [11] were improved) if the more general conditions (76), (77) replace (92)–(94), and

$$\bar{g}_0(r) \leqslant g_2(r), \quad r \in [0, R], \tag{105}$$

is satisfied. Moreover if strict inequality holds in (105) we obtain more precise error bounds. Indeed, define the sequence $\{\bar{r}_n\}$ $(n \geqslant 0)$, using (77), g_2 instead of (93), $\bar{g}_0$, respectively (with $\bar{g}_1 = g_1$, $\alpha = \bar{\alpha}$, $\beta = \bar{\beta}$), by

$$\begin{aligned} &\bar{r}_0 = r_0, \qquad \bar{r}_1 = r_1, \\ &\bar{r}_{n+1} - \bar{r}_n = \frac{u(\bar{r}_n) - u(\bar{r}_{n-1}) + (1 - g_2(\bar{r}_{n-1}) - \bar{\alpha})(\bar{r}_n - \bar{r}_{n-1})}{g(\bar{r}_n)} \quad (n \geqslant 1). \end{aligned} \tag{106}$$

It can easily be seen using induction on n (see also the proof of Proposition 1 that follows) that

$$\bar{r}_{n+1} - \bar{r}_n < r_{n+1} - r_n, \tag{107}$$

$$\bar{r}_n < r_n, \tag{108}$$

$$\bar{r}^* - \bar{r}_n \leqslant r^* - r_n, \quad \bar{r}^* = \lim_{n \to \infty} \bar{r}_n, \tag{109}$$

and

$$\bar{r}^* \leqslant r^*. \tag{110}$$

Furthermore condition (77) allows us more flexibility in choosing functions and constants.

Remark 5. Returning back to Newton's method (8) (see also (28)), the iterations corresponding to (97) and (106) are (13) and (17), respectively. Moreover condition (105) reduces to (19), and in case $\gamma_0 < \gamma_1$, estimates (22)–(24) hold.

Remark 6. Our error bounds (87), (88) are finer than the corresponding ones (102) and (103), respectively in many interesting cases.

Let us choose

$$\alpha = \bar{\alpha}, \quad \beta = \bar{\beta}, \quad g_0(r) = \bar{g}_0(r), \quad g_1(r) = g_2(r) = \bar{g}_1(r), \quad \text{and}$$
$$\bar{g}_3(r) = g_3(r) \quad \text{for all } r \in [0, R].$$

Then we can show

Proposition 1. *Under the hypotheses of Theorems* 3 *and* 4, *further assume*

$$s_1 < r_1. \tag{111}$$

Then, the following hold:

$$s_n < r_n \quad (n \geqslant 1), \tag{112}$$
$$s_{n+1} - s_n < r_{n+1} - r_n \quad (n \geqslant 0), \tag{113}$$
$$s^* - s_n \leqslant r^* - r_n \quad (n \geqslant 0), \tag{114}$$

and

$$s^* \leqslant r^*. \tag{115}$$

Proof. It suffices to show (112) and (113), since then (114) and (115), respectively, can easily follow. Inequality (112) holds for $n = 1$ by (111). By (82) and (97) we get in turn

$$\begin{aligned} s_2 - s_1 &= \frac{\int_0^1 \{g_1(s_0+\theta(s_1-s_0))\,d\theta - g_2(s_0)+\alpha\}(s_1-s_0)+\int_{s_0}^{s_1} g_3(\theta)\,d\theta}{1-\beta-g_0(s_1)} \\ &< \frac{\int_0^1 \{\bar{g}_1(r_0+\theta(r_1-r_0))\,d\theta - \bar{g}_2(r_0)+\bar{\alpha}\}(r_1-r_0)+\int_{r_0}^{r_1} \bar{g}_3(\theta)\,d\theta}{1-\bar{\beta}-\bar{g}_0(r_1)} \\ &= \frac{u(r_1)-u(r_0)+g(r_0)(r_1-r_0)}{1-\bar{\beta}-\bar{g}_0(r_1)} = \frac{u(r_1)}{g(r_1)} = r_2 - r_1. \end{aligned} \tag{116}$$

Assume

$$s_{k+1} < r_{k+1} \tag{117}$$

and

$$s_{k+1} - s_k < r_{k+1} - r_k \tag{118}$$

hold for all $k \leqslant n$.

Using (82), (88), and (118) we obtain

$$
\begin{aligned}
s_{k+2} - s_{k+1} &= \frac{\int_0^1 \{g_1[s_k+\theta(s_{k+1}-s_k)]\,d\theta - g_2(s_k)+\alpha\}(s_{k+1}-s_k) + \int_{s_k}^{s_{k+1}} g_3(\theta)\,d\theta}{1-\beta-g_0(s_{k+1})} \\
&< \frac{\int_0^1 \{\bar g_1[r_k+\theta(r_{k+1}-r_k)]\,d\theta - \bar g_2(r_k)+\bar\alpha\}(r_{k+1}-r_k) + \int_{r_k}^{r_{k+1}} \bar g_3(\theta)\,d\theta}{1-\bar\beta-\bar g_0(r_{k+1})} \\
&= \frac{u(r_{k+1}) - u(r_k) + g(r_k)(r_{k+1}-r_k)}{g(r_{k+1})} = \frac{u(r_{k+1})}{g(r_{k+1})} = r_{k+2} - r_{k+1}.
\end{aligned}
$$

That completes the proof of Proposition 1. □

In order for us to include a case where operator G is nontrivial, we consider the following example for Theorem 2 (or Theorem 3).

Example 5. Let $X = Y = C[0,1]$ the space of continuous functions on $[0,1]$ equipped with the sup-norm. Consider the integral equation on $\bar U(x_0, R/2)$ given by

$$x(t) = \int_0^1 k\bigl(t, s, x(s)\bigr)\,ds, \tag{119}$$

where the kernel $k(t,s,x(s))$ with $(t,s) \in [0,1] \times [0,1]$ is a nondifferentiable operator on $\bar U(x_0, R/2)$. Define operators F, G on $\bar U(x_0, R/2)$ by

$$F(x)(t) = Ix(t) \quad (I \text{ the identity operator}), \tag{120}$$

$$G(x)(t) = -\int_0^1 k\bigl(t, s, x(s)\bigr)\,ds. \tag{121}$$

Choose $x_0 = 0$, and assume there exists a constant $\theta_0 \in [0,1)$, a real function $\theta_1(t,s)$ such that

$$\bigl\|k(t,s,x) - k(t,s,y)\bigr\| \leqslant \theta_1(t,s)\|x-y\| \tag{122}$$

and

$$\sup_{t\in[0,1]} \int_0^1 \theta_1(t,s)\,ds \leqslant \theta_0 \tag{123}$$

for all $t, s \in [0,1]$, $x, y \in \bar U(x_0, R/2)$.

Moreover choose in Theorem 3: $r_0 = 0$, $y_0 = y_{-1}$, $A(x,y) = A(x) = I(x)$, I the identity operator on X, $g_0(r) = r$, $\alpha = \beta = 0$, $g_1(r) = g_2(r) = 0$, and $g_3(r) = \theta_0$ for all $x, y \in \bar U(x_0, R/2)$, $r, s \in [0,1]$ (similar choices for Theorem 3). It can easily be seen that the conditions of Theorem 2 hold if

$$t^* = \frac{\eta}{1-\theta_0} \leqslant \frac{R}{2}. \tag{124}$$

3. Local convergence of method (2)

In order to cover the local case, let us assume x^* is a zero of Eq. (1), $A(x^*, x^*)^{-1}$ exists and for any $x, y \in \bar{U}(x^*, r) \subseteq \bar{U}(x^*, R)$, $t \in [0, 1]$,

$$\left\| A(x^*, x^*)^{-1}\left[A(x, y) - A(x^*, x^*)\right]\right\| \leqslant \bar{h}_0\left(\|x - x^*\|, \|y - x^*\|\right) + \bar{a} \tag{125}$$

and

$$\begin{aligned}&\left\| A(x^*, x^*)^{-1}\left\{\left[F'\left(x^* + t(y - x^*)\right) - A(x, y)\right](y - x^*) + G(y) - G(x^*)\right\}\right\| \\ &\quad \leqslant \left[\bar{h}_1\left(\|y - x^*\|(1 + t)\right) - \bar{h}_2\left(\|y - x^*\|\right) + \bar{h}_3\left(\|x - x^*\|\right) + \bar{b}\right]\|y - x^*\|,\end{aligned} \tag{126}$$

where $\bar{h}_0, \bar{h}_1, \bar{h}_2, \bar{h}_3, \bar{a}, \bar{b}$ are as h_0, h_1, h_2, h_3, a, b, respectively. Then exactly as in (56) but using (125), (126), instead of (25), (26) we can show the following local result for method (2).

Theorem 5. *Assume that there exists a solution of equation*

$$f(\lambda) = 0 \tag{127}$$

in $[0, R]$, *where*

$$f(\lambda) = \int_0^1 \left[\bar{h}_1\left((1 + t)\lambda\right) - \bar{h}_2(\lambda)\right] dt + \bar{h}_3(\lambda) + \bar{h}_0(\lambda, \lambda) + \bar{a} + \bar{b} - 1. \tag{128}$$

Denote by λ_0 *the smallest of the solutions in* $[0, R]$. *Then, sequence* $\{x_n\}$ $(n \geqslant -1)$ *generated by Newton-like method* (2) *is well defined, remains in* $\bar{U}(x^*, \lambda_0)$ *for all* $n \geqslant 0$ *and converges to* x^* *provided that* $x_{-1}, x_0 \in \bar{U}(x^*, \lambda_0)$. *Moreover the following error bounds hold for all* $n \geqslant 0$:

$$\|x^* - x_{n+1}\| \leqslant p_n, \tag{129}$$

where

$$p_n = \frac{\{\int_0^1 [\bar{h}_1((1+t)\|x_n - x^*\|) - \bar{h}_2(\|x_n - x^*\|)]\, dt + \bar{a} + \bar{h}_3(\|x_{n-1} - x^*\|)\}}{1 - \bar{b} - \bar{h}_0(\|x_n - x^*\|)}\|x_n - x^*\|. \tag{130}$$

Application 3. Let us again consider Newton's method, i.e., $F'(x) = A(x, y)$, $G(x) = 0$, and assume

$$\left\| F'(x^*)^{-1}\left[F'(x) - F'(x^*)\right]\right\| \leqslant \lambda_1\|x - x^*\| \tag{131}$$

and

$$\left\| F'(x^*)^{-1}\left[F'(x) - F'(y)\right]\right\| \leqslant \lambda_2\|x - y\| \tag{132}$$

for all $x, y \in \bar{U}(x^*, r) \subseteq \bar{U}(x^*, R)$. Then we can set

$$\begin{aligned}&\bar{a} = \bar{b} = 0, \quad \bar{h}_3 = 0, \quad \bar{h}_1(r) = \bar{h}_2(r) = \lambda_2 r, \quad \text{and} \quad \bar{h}_0(r, r) = \lambda_1 r \\ &\quad \text{for all } r \in [0, R].\end{aligned} \tag{133}$$

Using (131), (132) we get

$$\lambda_0 = \frac{2}{2\lambda_1 + \lambda_2}. \tag{134}$$

Local results were not given in [11,12,21,23]. However Rheinboldt in [19] showed that under only (132) the convergence radius is given by

$$\lambda_3 = \frac{2}{3\lambda_2}. \tag{135}$$

But in general

$$\lambda_1 \leqslant \lambda_2. \tag{136}$$

Hence we conclude

$$\lambda_3 \leqslant \lambda_0. \tag{137}$$

The corresponding error bounds become

$$\|x_{n+1} - x^*\| \leqslant e_n, \tag{138}$$

$$\|x_{n+1} - x^*\| \leqslant e_n^1, \tag{139}$$

where

$$e_n = \frac{\lambda_2 \|x_n - x^*\|^2}{2[1 - \lambda_1 \|x_n - x^*\|]} \tag{140}$$

and

$$e_n^1 = \frac{\lambda_2 \|x_n - x^*\|^2}{2[1 - \lambda_2 \|x_n - x^*\|]}. \tag{141}$$

That is

$$e_n \leqslant e_n^1 \quad (n \geqslant 0). \tag{142}$$

If strict inequality holds in (136) then (137) and (142) hold as strict inequalities also (see also Example 4).

Remark 7. As noted in [1–9,25] the local results obtained here can be used for projection methods such as Arnoldi's, the generalized minimum residual method (GMRES), the generalized conjugate residual method (GCR), for combined Newton/finite projection methods and in connection with the mesh independence principle to develop the cheapest and most efficient mesh refinement strategies.

Remark 8. The local results can also be used to solve equations of the form $F(x) = 0$, where F' satisfies the autonomous differential equation [4,8,16]

$$F'(x) = P\big(F(x)\big), \tag{143}$$

where $P : Y \to X$ is a known continuous operator. Since $F'(x^*) = P(F(x^*)) = P(0)$, we can apply our results without actually knowing the solution x^* of Eq. (1).

Example 6. Let $X = Y = \mathbf{R}$, $\bar{U}(x^*, R) = \bar{U}(0, 1)$, $G = 0$, $A(x, y) = F'(x)$, and define function F on $\bar{U}(0, 1)$ by

$$F(x) = e^x - 1. \tag{144}$$

Then we can set $P(x) = x + 1$ in (143). Using (132) we get $\lambda_2 = e$. Moreover by (144) we get

$$\begin{aligned} F'(x) - F'(x^*) = e^x - 1 &= x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots \\ &= \left(1 + \frac{x}{2!} + \cdots + \frac{x^{n-1}}{n!} + \cdots\right)(x - x^*) \end{aligned} \tag{145}$$

and

$$F'(x^*)^{-1}\big[\big\|F'(x) - F'(x^*)\big\|\big] \leqslant (e-1)\|x - x^*\|.$$

That is $\lambda_1 = e - 1$. By (134) and (135) we get

$$\lambda_3 = 0.245252961 \tag{146}$$

and

$$\lambda_0 = 0.254028662. \tag{147}$$

That is our convergence radius λ_0 is larger than the corresponding one λ_3 due to Rheinboldt and our error bounds (140) are also finer than (141) so that (142) holds as a strict inequality. Finally note that all these improvements are made using the same hypotheses/information as in the earlier results. This observation is important in computational mathematics, since a wider choice of initial guesses x_0 becomes available (see also Remark 7).

The results obtained here can be extended to m-point methods ($m > 2$ an integer) [4–8, 17], and can be used in the solution of variational inequalities [23].

References

[1] F.L. Allgower, K. Böhmer, F.A. Potra, W.C. Rheinboldt, A mesh independence principle for operator equations and their discretizations, SIAM J. Numer. Anal. 23 (1986) 160–169.

[2] J. Appel, E. DePascale, P.P. Zabrejko, On the application of the Newton–Kantorovich method to nonlinear integral equations of Uryson type, Numer. Funct. Anal. Optim. 12 (1991) 271–283.

[3] I.K. Argyros, Improving the rate of convergence of Newton methods on Banach spaces with a convergence structure and applications, Appl. Math. Lett. 10 (1977) 21–28.

[4] I.K. Argyros, Advances in the Efficiency of Computational Methods and Applications, World Scientific, River Edge, NJ, 2000.

[5] I.K. Argyros, On the radius of convergence of Newton's method, Internat. J. Comput. Math. 77 (2001) 389–400.

[6] I.K. Argyros, A Newton–Kantorovich theorem for equations involving m-Fréchet-differentiable operators and applications in radiative transfer, J. Comput. Appl. Math. 131 (2001) 149–159.

[7] I.K. Argyros, F. Szidarovszky, Convergence of general iteration schemes, J. Math. Anal. Appl. 168 (1992) 42–52.

[8] I.K. Argyros, F. Szidarovszky, The Theory and Applications of Iteration Methods, CRC Press, Boca Raton, FL, 1993.

[9] P.N. Brown, A local convergence theory for combined inexact-Newton/finite-difference projection methods, SIAM J. Numer. Anal. 24 (1987) 407–434.
[10] E. Cătinas, On some iterative methods for solving nonlinear equations, Rev. Anal. Numér. Théor. Approx. 23 (1994) 47–53.
[11] X. Chen, T. Yamamoto, Convergence domains of certain iterative methods for solving nonlinear equations, Numer. Funct. Anal. Optim. 10 (1989) 37–48.
[12] J.E. Dennis, Toward a unified convergence theory for Newton-like methods, in: L.B. Rall (Ed.), Nonlinear Functional Analysis and Applications, Academic Press, New York, 1971, pp. 425–472.
[13] P. Deuflhard, G. Heindl, Affine invariant convergence theorems for Newton's method and extensions to related methods, SIAM J. Numer. Anal. 16 (1979) 1–10.
[14] P. Deuflhard, F.A. Potra, Asymptotic mesh independence of Newton–Galerkin methods via a refined Mysovskii theorem, SIAM J. Numer. Anal. 29 (1992) 1395–1412.
[15] J.M. Gutiérrez, M.A. Hernandez, M.A. Salanova, Accessibility of solutions by Newton's method, Internat. J. Comput. Math. 57 (1995) 239–247.
[16] L.V. Kantorovich, G.P. Akilov, Functional Analysis, Pergamon, Oxford, 1982.
[17] F.A. Potra, On an iterative algorithm of order 1.839 ... for solving nonlinear equations, Numer. Funct. Anal. Optim. 7 (1984–85) 75–106.
[18] F.A. Potra, V. Ptăk, Sharp error bounds for Newton's process, Numer. Math. 34 (1980) 67–72.
[19] W.C. Rheinboldt, An adaptive continuation process for solving systems of nonlinear equations, Banach Center Publ. 3 (1977) 129–142.
[20] J.W. Schmidt, H. Schwetlick, Ableitungsfreie Verfahren mit hohere Konvergenzgeschwindigkeit, Computing 3 (1968) 215–226.
[21] J.W. Schmidt, H. Leonhardt, Eingrenzung von Losungen mit der Regula Falsi, Computing 6 (1970) 318–329.
[22] J.W. Schmidt, Untere Fehlerschranken fur Regula-Falsi Verhafren, Period. Math. Hungar. 9 (1978) 241–247.
[23] L.U. Uko, J.O. Adeyeye, Generalized Newton-iterative methods for nonlinear operator equations, Nonlinear Studies 8 (2001) 465–477.
[24] T. Yamamoto, A convergence theorem for Newton-like methods in Banach spaces, Numer. Math. 51 (1987) 545–557.
[25] T.J. Ypma, Local convergence of inexact Newton methods, SIAM J. Numer. Anal. 21 (1984) 583–590.
[26] P.P. Zabrejko, D.F. Nguen, The majorant method in the theory of Newton–Kantorovich approximations and the Ptăk error estimates, Numer. Funct. Anal. Optim. 9 (1987) 671–684.